A Study on Vocal Training for Beginners by Analyzing Voice Pitch

So-Hyun Park¹, Aziz Nasridinov², Jeong-Yong Byun², Young-Ho Park¹,*

¹Department of Multimedia Science, Sookmyung Women’s University, Cheongpa-ro 47 gil 100,Yongsan-gu, Seoul, 140-742, Korea
{sohyun1218, yhpark}@sm.ac.kr

²School of Computer Engineering, Dongguk University at Gyeongju
123 Dongdaero, Gyeongju, Gyeongbuk, 780-714, Korea
{aziz, byunjy}@dongguk.ac.kr

*Corresponding Author

Abstract. The vocal training is useful for beginners to learn how to correctly sing, and for professional singers to tune their voice and expand their frequency range. In the paper, we propose a study on vocal training by analyzing voice pitch on android platform. Vocal training requires a vocal coach who is always besides you to correct your mistakes. However, beginner who wants to learn singing for fun or professional singer who practices his vocal technique everyday rather needs a short-hand assistance to control his or her singing. Thus, the vocal training is useful for beginners to learn how to correctly sing, and for professional singers to tune their voice and expand their frequency range. On the other hand, vocal training requires a vocal coach who is always besides you to correct your mistakes. However, beginner who wants to learn singing for fun or professional singer who practices his vocal technique everyday rather needs a short-hand assistance to control his or her singing.

In the paper, we propose a study on vocal training system by analyzing voice pitch on an android platform. The proposed method is divided into four parts: preprocessing, feature extraction, feature matching and result output, that enables to perform efficient voice recognition. In addition, we describe the various training options of the proposed method, including several statistical analyses in order to describe the user’s progress in later stages.

Keywords: vocal training, voice pitch, pitch analysis.

1 Introduction

Learning and practicing how to sing can be a confusing process for beginners and professional singers, as there are many vocal features to follow, such as voice pitch, frequency range, etc. Vocal training is a process to instruct singers on how to improve their singing skills, develop their vocal techniques, and prepare for the performance. Specifically, vocal training can help to identify where singers hit the right note, and where they were off. Thus, the vocal training is useful for beginners to learn how to correctly sing, and for professional singers to tune their voice and expand their frequency range. On the other hand, vocal training requires a vocal coach who is always besides you to correct your mistakes. However, beginner who wants to learn singing for fun or professional singer who practices his vocal technique everyday rather needs a short-hand assistance to control his or her singing.

In the paper, we propose a study on a vocal training system by analyzing voice pitch on an android platform. The proposed method is divided into four parts: preprocessing, feature extraction, feature matching and result output. In preprocessing step, we first capture the user’s vocal through smartphone microphone and digitize it.
by sampling. In feature extraction step, a number of acoustic parameters such as Mel Frequency Cepstral Coefficient, MFCC, are extracted from the user’s voice, and pitch detection is performed. In the feature matching step, we compare the note values and notice the user about the error in pitch when he or she sings a wrong note. In addition, we describe the various training options of the proposed method, including several statistical analyses in order to describe the user’s progress in later stages.

The rest of the paper proceeds as follows. Section 2 discusses the related work. Section 3 describes the proposed method. Section 4 highlights conclusions.

2 Related Work

There have been a number of research and commercial applications focused on vocal training. We briefly review each of these approaches in this section. Mendes et al. [1] investigated the effects of vocal training on the parameters of the singing voice. Specifically, a group of college voice majors were examined during four semesters. Acoustic measures included Maximum Phonational Frequency Range, MPFR, vibrato and the presence of the singer’s formant. The result of the research reveals that four semesters of Vocal Training, VT, had a significant effect on the MPFR while effect on vibrato and the singer’s formant is not observed. Similarly, LeBorgne and Weinrich [2] studied the effects of VT for singers over a period of nine months. Vocal music students, who were involved in an intensive vocal performance curriculum, participated in this study. The results of this study demonstrate that training expands the singer's ability to increase frequency range and improve dynamic control. Pandiaraj et al. [3] proposed a proficient vocal training system. The main idea of the proposed roach is that it detects pitch from a vocal of both amateur and professional vocalist, compares, perceives error in the pitch, and displays inaccuracy of the amateur’s voice. Numerous researches are focused on developing vocal training systems.

Recently, it became comfortable for a user to train his/her voice via smartphone. A number of commercial applications are developed in order to perform voice training on smartphone. Vocalist is a smartphone application that develops your singing skills and vocal range by providing various singing exercises. The singing exercises play a short melody that has to be sung by a user. Vocalist shows you instant feedback on where you sung the right note, and where you made a mistake. Vocalist has a problem which is not including efficient music education curriculum. Voice Training is another application that assists you to develop singing skills. Similar to Vocalist, it plays vocal exercises as one to five notes, you can listen the following notes and sing the same notes. Voice Training has a problem which is not including functions that recognize user’s voice and match feature in order to give feedback.

3 Proposed Method of Analyzing Voice Pitch

In this section, we describe the proposed method. Fig.1. demonstrates overall process of the proposed method. The proposed method is divided into four parts as follows to
recognize user’s voice, analyze pitch and identify accepted voice’s pitch is correct or not.

- **Step 1: Preprocessing.** The user of the proposed system uses the smartphone application that is developed on android platform, where he or she is asked to repeat the given notes. In the preprocessing step, we first capture the user’s voice through smartphone microphone, and save it in SD card of the device. We implemented MediaRecorder named in the paper, which recognize the human voice, in order to perform this action. Once the user’s voice is captured, we sample the analog input into a series of 16 bit values at a particular sampling frequency, which is ranging from 8 to 22 KHz that smartphone microphone can recognize [7].

- **Step 2: Feature Extraction.** In this step, a number of acoustic parameters such as the MFCC are extracted from the user’s voice, and pitch detection is performed. MFCC is used to determine the pitch of the signal and is calculated using the following equation [4, 6]:

\[
    c(\tau) = F\{log(|F\{x[n]\}|^2)}
\]

Where \( F \) indicates Fourier Transform, and \( x[n] \) is the signal in the time domain. \(|F\{x[n]\}|^2\) is the power spectrum estimate of the signal. The power cepstrum shows the peaks corresponding to the period of the pitch frequency in the spectrum, and hence performs pitch detection.

- **Step 3: Matching Features.** Once the acoustic features are extracted, we begin matching those features with a set of standard pitch frequencies that is stored in database. This enables to determine whether the user hit the right note or made a mistake [5].

- **Step 4: Result Output.** The previous step provides us with detailed information of where user made a mistake. According to this information, we define a matching rate, calculate the score and output it to the user. Matching Features step also enables us to know more detailed information on which note the user made a mistake. This enables us to provide the user with the correction instructions.

![Fig.1. Overall Flow of Proposed Method.](image)

The proposed method can have the following training options in the future work. Lessons that are categorized into several difficulty levels ranging from beginner to master class. Scientific description of the singing results explains where the user made a mistake in detail. Specifically, we graphically demonstrate the score of singing during all period of training, as well as progress made by the user during this period.
Correct instructions provide users with tips to improve his or her singing technique. Karaoke mode enables you to practice singing the real songs. Social networks among friends enable to share their progress with his or her friends. This can make learning period more enjoyable. The proposed system keeps the profile of all users singing data and performs various statistical analyses in order to show the user’s progress in later stages. We use correlation analysis, mean analysis and basic data mining based prediction. Through the future work, we will show the detail algorithms, analysis, and the experimental results for the voice pitch system simply sketched.

4 Conclusion

In the paper, we have proposed a study on vocal training system by analyzing voice pitch on android platform. The proposed method is useful for beginners to learn how to sing, for professional singers to practice and improve their singing skills, for vocal coaches to simplify teaching process and save their time. It can save from the costly service of vocal coaches as well as replace them in cases where they are not needed. In the future, we will describe the performance evaluation of the proposed method and concentrate on measuring the accuracy of feature matching, the increase of user’s vocal frequency range.

Acknowledgement. This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012003797).

References